1. **训练RSPrompter**

代码见RSPrompter

**环境安装：**见github官网（有点复杂，有些库比较大，失败再重来，多次之后还是可以成功的。技巧：有些库找不到或装不上，url超时，在pip 命令后加上-i http://pypi.douban.com/simple/ --trusted-host pypi.douban.com）

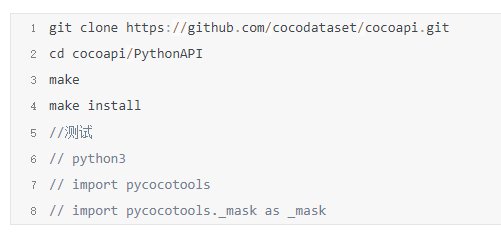
安装好环境后，tools/train.py跑起来，选择不展示训练结果的模式下，一直在进行数据处理，没进入训练迭代，显卡利用率一直为0；选择展示训练结果的模式，也是一样的表现，看不到任何展示结果---

1. **直接用coco\_seg的数据集训练SAM, 可直接输出mask, 通过mask可以计算得到bbox， 后续可以用输出的bbox与提取的特征相乘，用一个分类模块对bbox进行分类，即可得到手术器械的instance\_seg;**

代码见lightning\_sam/segment\_anything/my\_train.py

**环境安装：**

1. 安装pycocotools:



1. 安装segment\_anything: pip install 'git+https://github.com/facebookresearch/segment-anything.git'
2. 安装python-box: pip install python-box

训练时，可以通过config配置要finetune的模块，比如可以冻结image\_encoder和prompt\_encoder，只训练mask\_decode；只训练mask\_decode时，模型不收敛，怀疑是训练集图片质量太差，编码器不适应，所以同时finetune image\_encoder和mask\_decoder，但模型还是不收敛；

不管损失函数是原来的20. \* loss\_focal + loss\_dice + loss\_iou， 还是loss\_dice + loss\_iou或者只用loss\_dice，模型都不收敛；

用lightning\_sam/segment\_anything/my\_predict.py对模型进行测试时，原本的SAM对每张测试图片输出的mask个数基本都大于5， 而finetune出来的模型，输出的mask都为空；

综上，用coco\_seg的数据集finetune SAM失败，后续的工作没法继续；分析可能的原因：1）该分割任务比较困难（首先是类别不平衡（差距100来倍），然后是图片质量差（视频里采样出来的，会有模糊，失焦的问题），类别分类困难，每个工具的特征比较相似（同时镜头视野又比较小，每个工具在每一张图中只能够展现出部分特征），2）SAM是针对交互式分割的场景提出的，需要prompt, 没有prompt的场景不适合；

1. **思路：用全自动分割模型比如unet进行粗分割，将unet分割得到的mask作为prompt用来提示SAM, 组合起来就可以通过SAM得到全自动的手术器械分割mask；**

具体：直接用coco\_seg的数据集训练unet；

代码见lightning\_sam/segment\_anything/train\_unet.py（在lightning\_sam下，环境也一样）

损失函数直接用dice\_loss， 训练过程中感觉loss有下降趋势，但是每个epoch结束，平均loss总是只有轻微波动，模型也时无法收敛；

分析原因：除了图片质量差之外，图像特征非常多样化（即不同video的图像特征差异非常大），分割目标识别非常困难；